The Report committee for Aidyn Beisenov

Certifies that this is the approved version of the following report:

The Life Insurer Risk-Based Capital Ratio: Panel Data Analysis

APPROVED BY

SUPERVISING COMMITTEE:

Supervisor: ______________________________
Thomas W. Sager

____
Daniel A. Powers

The Life Insurer Risk-Based Capital Ratio: Panel Data Analysis

by

Aidyn Beisenov, B.S.

Report

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Science in Statistics

The University of Texas at Austin

May 2013

Acknowledgements

I would like to express my sincere gratitude to Dr. Thomas Sager for his enthusiasm, guidance and advice throughout the writing of this report.

My sincere thanks also go to Dr. Daniel Powers for valuable comments and support during my graduate study.

May 2013

Abstract

The Life Insurer Risk-Based Capital Ratio: Panel Data Analysis

Aidyn Beisenov, M.S.Stat.

The University of Texas at Austin, 2013

Supervisor: Thomas W. Sager

Many studies suggest the ability of the NAIC Risk-Based Capital ratio (RBC ratio) to predict insurer insolvency. Based on the US life insurer (insurer) data for the period of 2005 to 2008, this study finds explanatory variables that have a statistically significant relationship with the RBC ratio. Advantages of panel data over cross-sectional and time series data analysis are exploited to make valid inference on coefficients of the explanatory variables. Testing for unobserved insurer and time effects and for dependence between these effects and the explanatory variables indicates the appropriateness of the fixed insurer and time effects model. Based on the ordinary least squares estimates, it is found that insurers' size, capital-to-asset ratio, and return on capital have a statistically significant relationship with the RBC ratio. Additionally, health product, annuity product, opportunity, and regulatory risks of insurers are related to the RBC ratio. Accounting for heteroscedasticity and autocorrelation for a given insurer yields the same coefficient estimates, but increased standard errors.

Table of Contents

List of Tables

Chapter 1. Introduction

1.1. The Risk-Based Capital ratio

The insurance industry in the US is under the regulation of state insurance commissioners, since various stakeholders will be affected if their insurers fail to meet obligations. There are few insolvency cases in every year, since it is very unlikely that the regulator will not intervene in an insurer with poor financial performance before it becomes insolvent. To identify problematic insurers, the regulators utilize the risk-based capital ratio (RBC ratio), a measure of the adequacy of an insurer's capital relative to its risk of insolvency (Grace et al. 1998). The regulators are assisted by the National Association of Insurance Commissioners (NAIC) to establish standards and best practices, conduct peer review, and coordinate their regulatory oversight (NAIC 2009). Insurers in the US file annual and quarterly financial statements, and the NAIC uses them to maintain and update the Financial Data Repository.

> The NAIC employs the risk-based capital (RBC) system that has two main components: 1) the risk-based capital formula, that establishes a hypothetical minimum capital level that is compared to a company's actual capital level, and 2) a risk-based capital model law that grants automatic authority to the state insurance regulator to take specific actions based on the level of impairment (NAIC 2009).

The authorized control level RBC for an insurer is annually calculated and defined by the NAIC as the amount of required capital that the insurance company must maintain based on the inherent risks in the insurer's operations. Separate RBC formulas that take into account size and risk profile exist for life, health, and property/casualty insurers (NAIC 2009).

An insurer's RBC ratio is defined as 50 times the ratio of an insurer's total adjusted capital to its authorized control level RBC. Basically, the RBC ratio determines if an insurer sustains capital that adequately meets its risk exposure. There are 5 levels of actions that the regulators can take based on the RBC ratios.

1. No action is taken in the case the RBC ratio greater than 200. This means an insurer's capital is deemed sufficient to support its operations and investments.
2. Company action level: If the RBC ratio is between 150 and 200, an insurer is required to file a special report to the regulator. In the report the insurer must outline comprehensive financial plan. The failure to submit the report triggers the Regulatory action level.
3. Regulatory action level: The RBC ratio of between 100 and 150 not only requires the insurer to file the report, but the regulator to perform necessary examinations or analyses. It may result in corrective order for the insurer to address the financial problems.
4. Authorized control level: If the RBC ratio is from 70 to 100, the law permits the regulator to take control of the insurer in addition to powers provided at the higher action levels.
5. Mandatory control level: The RBC ratio is less than 70, and the regulator is required to take the insurer under control.

Additionally, there is a trend test targeted on insurers with RBC ratios between 200 and 250. A negative trend below a certain level triggers the company action level (NAIC 2009).

The ability of the RBC ratio to predict insurer insolvency has been studied by researchers. Cummins et al. (1995) concluded that RBC ratios generally were significantly different for failed and surviving firms, but less than half of the firms that failed had RBC ratios corresponding to regulatory and company action levels in the property-liability insurance. Grace et al. (1998) found limited support for the hypothesis that the use of RBC ratios produced new information concerning insurer insolvency risk. We intend to make inference on relationships between a set of predictors and the RBC ratio of life insurers. Based on relatively new data, this study may reveal variables related to insolvency in the life insurance industry.

Reviewing studies by Baranoff and Sager (2002, 2011) on capital and risks in the US life insurance industry gave insight to conduct our study. Among other control variables, Baranoff and Sager use the risk-based capital ratio as a proxy for the force of regulatory pressure. To make valid inference on predictors, we conduct panel data analysis.

1. 2. Panel data

Panel data consist of units that are observed across time. Units can be individuals, households, firms, countries, etc. If there are N units, and each unit is measured over T time periods, the total number of observations will be NT. We can distinguish between balanced and unbalanced panel data. If panel data is balanced, all units include measurements in all time periods. In unbalanced panel data, all units in a data set have different numbers of observations due to missing values, and the total number of observations is not NT.

The analysis of panel data allows the model builder to learn about economic processes while accounting for heterogeneity across cross-sectional units as well as for dynamic effects that are not detectable in cross sections (Greene 2010, 343). The advantages of panel data over traditional cross-sectional and time-series data sets are central to our study.

According to Baltagi (1995, 2), virtually every graduate text in econometrics contains a chapter or a major section on the econometrics of panel data. The growth of applied studies and the methodological development of new econometric tools of panel data have been simply phenomenal since the seminal paper of Balestra and Nerlove in 1966 (Hsiao 2006). Hsiao claims, in 1986, there were 29 studies listing the key words: "panel data or longitudinal data", according to Social Sciences Citation index. By 2004, there were 687 and by 2005, there were 773 (Hsiao 2006). We updated information about a proliferation of panel data studies in recent years. In 2013, Social Sciences Citation index databases showed 2491 results where a title of scientific papers included the key word "panel data."

According to Hsiao, the growth of the number of panel data studies can be explained by three main factors: (i) data availability, (ii) greater capacity for modeling the complexity of human behavior than a single cross-section or time series data, and (iii) challenging methodology (Hsiao 2006).

The following are the most important of the advantages of panel data according to Hsiao (2006):

1. Panel data usually provides more degrees of freedom and reduces multicollinearity which allows more practical inference about the model parameters.

2. More sophisticated behavioral hypotheses can be tested. Panel data may allow controlling the effects of omitted variables given the intertemporal dynamics and the heterogeneity of the units. One can make more accurate inference not only on dynamic relationships, but also on predictions of individual outcomes given the pooled data.

3. Panel data usually consists of cross-sectional and time dimensions, and it may simplify inference on non-stationary time series and measurement errors.

Chapter 2

2.1. Panel data regression analysis

Regression analysis of panel data is the statistical technique considered in this report. In the following linear panel data regression, a double subscript is used on variables.

$Y_{it} = \beta_0 + X_{it}*\beta + u_{it}$,

$i = 1,\ldots, N$ and $t = 1,\ldots, T$, where i denotes units, and t denotes time periods. β_0 is the intercept, β is a M dimensional vector of slopes and X_{it} is the *it*-th observation on M explanatory variables. If the u_{it} are assumed iid $(0, \sigma_u^2)$, ordinary least squares (OLS) estimation applies to the model. This is called a pooled model, which will be considered in our study.

In regression analysis, a fundamental problem is omitted-variables bias (Greene 2010, 346). This bias may be mitigated by taking into account unobserved heterogeneity. The fixed-effects and random-effects models will be applied to control for unobserved heterogeneity. In these models, the error term is modeled as $u_{it} = \alpha_i + \varepsilon_{it}$. α_i is time-invariant, and is accounts for omitted variables and unobserved unit specific effect. ε_{it} denotes the regression error term, and, in the simplest case, they are assumed iid $(0, \sigma_\varepsilon^2)$. If unobserved time specific effects exist in addition to unit specific effects, the error term can also be modeled as $u_{it} = \alpha_i + \tau_t + \varepsilon_{it}$, where τ_t denotes the time specific effect. Hsiao refers to u_{it} as an incidental parameter, because when units, N and time series observations, T increase, so does the dimension of u_{it}. A general principle of obtaining valid inference of β in the presence of incidental parameters u_{it} is to find a proper transformation to eliminate u_{it} from the specification (Hsiao 2006).

A fundamental difference between the fixed and random effects models is the role of the unobserved heterogeneity. For example, given only unit specific effects, different assumptions about the correlation between the explanatory variables, X_{it}, and unit specific constant, α_i, lead to alternative specifications of regression models. If α_i is assumed to be correlated with X_{it}, the fixed effects model is accepted. In the fixed effects model α_i is a part of the intercept.

$$Y_{it} = \alpha_i + X_{it}*\beta + \varepsilon_{it},\ \varepsilon_{it} \sim iid(0, \sigma_\varepsilon^2)$$

In the case of the random effects model, α_i is assumed to be independent of X_{it} and act as a part of the regression error term.

$$Y_{it} = \beta_0 + X_{it}*\beta + \alpha_i + \varepsilon_{it},\ \alpha_i \sim iid(0, \sigma_\alpha^2),\ \varepsilon_{it} \sim iid(0, \sigma_\varepsilon^2)$$

α_i acts as an omitted variable that has been added to the model, and it affects the parameters in the model if correlated with the explanatory variables. If it is not correlated with the explanatory variables, the coefficients of the explanatory variables are not affected.

We introduce two approaches for estimating the fixed-effects model. The least squares dummy variable model (LSDV) fits the fixed effects model using dummy variables. LSDV is a well-known approach, since its estimates are relatively easy to interpret. Dummy variables account for the effects of those omitted variables that are specific to individual cross-sectional units but stay constant over time, and the effects that are specific to each time period but are the same for all cross-sectional units (Hsiao

1985, 29). The coefficients of dummy variables are parameters whose number increases as the number of units or time periods increases. This can lead to so-called incidental parameter problem. LSDV suffers from a substantial loss of degrees of freedom, and estimating (N-1) extra parameters may aggravate the problem of multicollinearity among regressors (Baltagi 1995, 11).

If the number of units or time periods is very large, we may want to use the within group effect model that does not involve dummy variables. The within group estimators subtracts unit means from observations and fits the regression model on data. The deviation from means transformation of variables eliminates observed and unobserved time-invariant heterogeneity. Its disadvantage is removing time-invariant explanatory variables from the model, so the effects of those variables cannot be estimated.

2.2. Data processing

Data used in this study is obtained from the NAIC life insurers RBC data set provided by Professor Thomas W. Sager. Life insurers' annual statements constitute the NAIC data set which contains cross-sections of life insurers over years 1993-2008.

The data set is arranged in long form, so we do not need to rearrange it. Initially there were 18056 records on insurers' RBC ratios. After removing insurers with RBC ratios which are negative or greater than 10000, the number of RBC ratio records reduced from 18506 to 17400.The negative and extremely large RBC ratios are possibly due to accounting anomalies and very small value of authorized capital (Baranoff and Sager 2004). Since this study focuses on the years from 2005 to 2008, we removed all the records not corresponding to these years from the data set, which reduced the number of the RBC ratio records to 3317. The original data set was unbalanced, and it required us to delete observations with missing data. We need at least 2 observations per insurer in order to build panel data models such as a fixed effects model. Removing insurers with less than 2 observations with non-missing values resulted in sample data that included 862 US life insurers.

2.3. Explanatory variables

The goal of this study is to identify the variables that account for variation in the RBC ratio of life insurers. Life insurers' financial statements provide a great number of variables. The selection of explanatory variables is guided by the results of previous studies that have been conducted on life insurers' data. Baranoff and Sager (2011) was the main source that provided insight to establish a pool of explanatory variables. Our goal is to examine the significance of those variables, rather than to identify new variables. The variables are defined as follows.

The RBC ratio is the dependent variable: 50 times an insurer's total adjusted capital divided by its authorized control level RBC in the natural logarithm form. Total adjusted capital is difference between assets and liabilities. (LogRBCratio).

Size of an insurer: the natural logarithm of an insurer's total assets is used as a proxy (LogATotal);

Capital ratio: the book value of an insurer's capital over total assets (CAP);

A return on capital: A ratio of income to market capital (RetOnCap);

Annuity product risk: annuity writings divided by total writings (ProdARisk);

Health product risk: health writings divided by total writings (ProdHRisk);

Reinsurance: reinsurance writings divided by total writings (preinsur);

Regulatory asset risk: regulatory asset risk divided by total invested assets (pRegARisk);

Opportunity asset risk: opportunity asset risk divided by total invested assets (pOppARisk);

An indicator variable for the governance structure (if stock, Ntype=1, and if mutual, Nype=0);

An indicator variable for whether or not the insurer is a member of a group of affiliated companies (if member, Ngroup=1, and if non-member, Ngroup=0).

The statistical software package SAS was used to obtain the descriptive statistics and perform regressions.

We intend to compare different estimation methods as well as different models.

Chapter 3

3.1. Exploratory data analysis

The table 1 presents descriptive statistics of the variables observed during a period of 2005 to 2008. The mean, standard deviation, minimum and maximum values for all the variables are provided in the table.

Variable	Label	N	Mean	Std Dev	Min	Max
group	NAIC group number	3317	661.05	966.06	0	4669.00
Preinsur	Reinsurance writings / Total writings	3297	0.14	0.31	-4.34	1.22
Ntype	Org type (1=stock)	3317	0.92	0.26	0	1.00
Ngroup	Indicator for member of group (1=yes)	3317	0.77	0.41	0	1.00
ProdHRisk	Health writings / Total writings	3297	0.29	0.37	-0.12	1.76
ProdARisk	Annuity writings / Total writings	3297	0.17	0.30	-0.008	1.01
RBCratio	100*Mkt cap / (2*Auth cap)	3317	886.27	1275.40	4.87	9949.48
RetOnCap	Income / Mkt capital	3317	0.01	1.10	-58.25	9.38
CAP	Mkt capital / Total assets	3317	0.31	0.27	0.002	1.82
LogATotal	log(Total assets)	3317	19.15	2.76	11.70	26.41
pRegARisk	Regulatory asset risk / Total invested assets	3317	0.02	0.04	0.00004	0.29
pOppARisk	Opportunity asset risk / Total invested assets	3317	0.003	0.004	0.00009	0.06

Table 1: Descriptive statistics for the variables (all years combined: 2005-2008)

The table shows that the number of observations is not the same for all the variables. Approximately, 92% of the insurers are stock companies, and 77% are members of groups of affiliated companies. On average, health, annuity, and reinsurance writings account for 29%, 17%, and 14% of total writings, respectively. It implies that 40% of total writings are life writings which are excluded from the set of predictors to avoid perfect collinearity.

A pairwise correlation analysis of the predictors is conducted. Pearson correlation coefficients measure a linear relationship between any pair of the variables. A cut-off value 0.8 is used to identify highly correlated variables. The Pearson correlation coefficient matrix reveals that no harmful pairwise correlation exists (see Appendix A).

We fit separate multiple regression models for a cross-section of the insurers in each year of the period of 2005 to 2008. Parameters in the models are estimated by means of ordinary least squares, and predictors that consistently have a statistically significant relationship with the RBC ratio are of interest.

We obtained separate plots of residuals against predicted values of the RBC ratio for every given year. In Appendix B, a funnel-shaped pattern on residual plots indicates increasing error variance.

We perform the natural logarithm transformation of the response variable in order to capture the error heteroscedasticity and possible nonlinearity of regression. After fitting multiple regression models for each year, plots of residuals against predicted values of the response variable show random scatter plots about 0 (see Appendix C). This indicates that the errors terms are homoscedastic.

Assessment of normality requires goodness-of-fit tests such as Shapiro-Wilk, Kolmogorov-Smirnov, Cramer-Von Mises, and Anderson-Darling. The null hypothesis is the values of the residuals are a random sample from the normal distribution. The tests are performed for every given year, and the p-values allow rejecting the null hypotheses at the 0.05 significance level. Since the sample sizes are large, even small departures from normality may produce highly significant test results. Thus, we obtain histograms and normal probability plots of residuals to make a final decision. The plots in Appendix D seem to approximately indicate normality.

The outputs for multiple regressions including the variance inflation factor (VIF) are presented in the Table 2. The VIF measures the inflation in the variances of the parameter estimates caused by multicollinearity. A cut-off value 10 is used, and the results suggest no harmful multicollinearity.

The log of total assets seems not to have significant relationship with the log of RBC ratio in any given year. The coefficient estimates vary from negative to positive throughout the period. Capital to asset ratio has a uniformly significant relationship with the response variable, and the coefficient estimates and standard errors are similar. Income to capital ratio has a non-significant and negative coefficient estimate only in 2006. Health Risk is highly significant in all years, and its coefficient estimates and standard errors are analogous. Annuity risk is significant only in 2007, and reinsurance is non-significant. Regulatory asset risk is uniformly significant, but the parameter estimates are not quite different. Opportunity asset risk is non-significant in all years, and parameter estimates are similar, except in 2005. The indicator variables for group membership and company structure are uniformly non-significant. We conclude that cross-sectional study of the insurer data may lead to different estimates and decisions on parameters in different years. This can be due to bias caused by omitted variables, thus panel data analysis is conducted to investigate relationships between RBC ratio and predictors while controlling for the unobserved insurer and time specific effects.

Year	Variable	Estimate	Standard error	T value	p-value	VIF
2005	Intercept	6.3048	0.2624	24.03	<0.0001	
	LogATotal	-0.0126	0.0124	-1.01	0.3131	2.37
	CAP	2.0867	0.1268	16.46	<0.0001	2.35
	RetOnCap	0.3033	0.0835	3.63	0.0003	1.04
	ProdHRisk	-0.8903	0.0732	-12.16	<0.0001	1.47
	ProdARisk	0.0377	0.094	0.4	0.6889	1.63
	Preinsur	-0.2402	0.0834	-2.88	0.0041	1.31
	pRegARisk	-5.301	1.2582	-4.22	<0.0001	6.08
	pOppARisk	-18.1029	21.6251	-0.84	0.4028	7.03
	Ngroup	0.075	0.0587	1.28	0.2016	1.23
	Ntype	-0.0165	0.0844	-0.2	0.8448	1.05
2006	Intercept	5.968	0.2706	22.05	<0.0001	
	LogATotal	0.0009	0.0126	0.08	0.9383	2.21
	CAP	2.106	0.1267	16.62	<0.0001	2.06
	RetOnCap	-0.057	0.0467	-1.22	0.2222	1.00
	ProdHRisk	-0.8308	0.0744	-11.17	<0.0001	1.36
	ProdARisk	0.1664	0.1003	1.66	0.0976	1.64
	Preinsur	0.0879	0.0767	-1.14	0.2526	1.24
	pRegARisk	-6.6483	1.3224	-5.03	<0.0001	6.02
	pOppARisk	12.806	28.9742	0.44	0.6586	6.24
	Ngroup	0.1065	0.0624	1.71	0.0882	1.21
	Ntype	-0.0405	0.0895	-0.45	0.6513	1.05
2007	Intercept	6.0144	0.2836	21.21	<0.0001	
	LogATotal	-0.0014	0.0134	-0.1	0.9181	2.33
	CAP	2.1142	0.1319	16.03	<0.0001	2.14
	RetOnCap	0.063	0.0116	5.43	<0.0001	1.01
	ProdHRisk	-0.7977	0.0776	-10.28	<0.0001	1.44
	ProdARisk	0.2351	0.103	2.28	0.0227	1.63
	Preinsur	-0.2679	0.089	-3.00	0.0028	1.3
	pRegARisk	-8.204	1.4653	-5.6	<0.0001	6.82
	pOppARisk	14.0613	16.279	0.86	0.388	7.59
	Ngroup	0.041	0.064	0.64	0.5225	1.23
	Ntype	0.0295	0.0935	0.32	0.7524	1.05
2008	Intercept	5.75	0.29	19.81	<0.0001	
	LogATotal	0.012	0.014	0.87	0.3857	2.36
	CAP	2.2632	0.1319	17.15	<0.0001	2.12
	RetOnCap	0.282	0.0444	6.35	<0.0001	1.07
	ProdHRisk	-0.9695	0.0788	-12.3	<0.0001	1.42
	ProdARisk	0.056	0.1034	0.54	0.5886	1.7
	Preinsur	-0.389	0.0925	-4.21	<0.0001	1.29
	pRegARisk	-8.6719	2.0734	-4.18	<0.0001	9.82
	pOppARisk	10.7735	10.265	1.05	0.2943	9.33
	Ngroup	0.0028	0.0655	0.04	0.9663	1.23
	Ntype	0.0713	0.0961	0.74	0.4581	1.06

Table 2: Multiple regressions results (response is LogRBCratio)

3.2. A pooled model

A pooled regression model, which assumes a constant intercept and slope for all the insurers, is considered first. It is basically an ordinary least squares regression model.

$LogRBCratio_{it} = \beta_0 + \beta_1 * LogATotal_{it} + \beta_2 * CAP_{it} + \beta_3 * RetOnCap_{it} + \beta_4 * ProdARisk_{it} + \beta_5 * ProdHRisk_{it} + \beta_6 * Preinsur_{it} + \beta_7 * pRegARisk_{it} + \beta_8 * pOppARisk_{it} + \beta_9 * Ngroup_{it} + \beta_{10} * Ntype_{it} + \varepsilon_{it}$.

In the proposed regression equation, β_0 is the intercept, β_i is a slope of the *i*-th explanatory variable, and ε_{it} is the error term.

Model Description	
Estimation Method	Pooled
Number of Cross Sections	862
Time Series Length	4

Table 3A

Fit Statistics			
SSE	1475.1144	**DFE**	3217
MSE	0.4585	**Root MSE**	0.6772
R-Square	0.4010		

Table 3B

Parameter Estimates						
Variable	**DF**	**Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**Label**
Intercept	1	6.018891	0.1370	43.93	<.0001	Intercept
LogATotal	1	-0.00108	0.00647	-0.17	0.8670	log(Total assets)
CAP	1	2.144223	0.0638	33.59	<.0001	Mkt capital / Total assets
RetOnCap	1	0.073528	0.0107	6.90	<.0001	Income / Mkt capital
ProdHRisk	1	-0.84696	0.0382	-22.20	<.0001	Health writings / Total writings
ProdARisk	1	0.116486	0.0506	2.30	0.0213	Annuity writings / Total writings
Preinsur	1	-0.24586	0.0425	-5.78	<.0001	Reinsurance writings / Total writings
pRegARisk	1	-6.24044	0.4390	-14.21	<.0001	Regulatory asset risk / Total invested assets
pOppARisk	1	-2.57298	3.8010	-0.68	0.4985	Opportunity asset risk / Total invested assets
Ngroup	1	0.0644	0.0314	2.05	0.0406	Indicator for member of group (1=yes)
Ntype	1	0.011922	0.0456	0.26	0.7938	Org type (1=stock)

Table 3C: Proc Panel output of the pooled model

R-Squared is equal to 0.40 which means the pooled model accounts for 40.1 % of the total variance in the response variable. Size, opportunity asset risk, and governance type turn out to be variables that do not have statistically significant relationships with the response variable.

This model is considered a naive model, since the RBC ratio may be affected by unobserved insurer heterogeneity and time-specific effects in addition to the explanatory variables. Disregarding control for these effects may cause invalid inference on the model parameters. We estimate a fixed-effects as well as random-effects models and obtain tests for presence of fixed or random effects.

3.3. A fixed-effects model

There are several approaches to fit a fixed-effects model. One of them is the least squares dummy variable method which we can employ using PROC PANEL in SAS. To avoid perfect collinearity between parameters PROC PANEL drops one dummy variable from a model, but a parameter for the variable is estimated as the intercept of the model.

To control for and test the time-invariant heterogeneity among insurers, we estimate a one-way fixed firm effects model. We fit a multivariate regression model with fixed firm effects, $LogRBCratio_{it} = \beta_0 + \beta_1*LogATotal_{it} + \beta_2*CAP_{it} + \beta_3*RetOnCap_{it} + \beta_4*ProdARisk_{it} + \beta_5*ProdHRisk_{it} + \beta_6*Preinsur_{it} + \beta_7*pRegARisk_{it} + \beta_8*pOppARisk_{it} + \alpha_i + \varepsilon_{it}$, where the α_i are dummy variables for firm effects. It includes dummy variables for each insurer except the last one, which becomes the intercept.

Model Description	
Estimation Method	FixOne
Number of Cross Sections	862
Time Series Length	4

Table 4A

Fit Statistics			
SSE	209.9096	**DFE**	2356
MSE	0.0891	**Root MSE**	0.2985
R-Square	0.9148		

Table 4B

F Test for No Fixed Effects			
Num DF	**Den DF**	**F Value**	**Pr > F**
861	2356	16.49	<.0001

Table 4C

Variable	DF	Estimate	Standard Error	t Value	Pr > |t|	Label
LogATotal	1	0.190349	0.0263	7.25	<.0001	log(Total assets)
CAP	1	2.666397	0.0931	28.65	<.0001	Mkt capital / Total assets
RetOnCap	1	0.049068	0.00629	7.80	<.0001	Income / Mkt capital
ProdHRisk	1	-0.36288	0.0990	-3.67	0.0003	Health writings / Total writings
ProdARisk	1	-0.31828	0.1130	-2.82	0.0049	Annuity writings / Total writings
Preinsur	1	0.011783	0.0619	0.19	0.8492	Reinsurance writings / Total writings
pRegARisk	1	-8.31211	0.5491	-15.14	<.0001	Regulatory asset risk / Total invested assets
pOppARisk	1	-0.88275	1.8900	-0.47	0.6405	Opportunity asset risk / Total invested assets
Ngroup	1	0.004517	0.0636	0.07	0.9434	Indicator for member of group (1=yes)
Ntype	1	0.096805	0.2446	0.40	0.6923	Org type (1=stock)

Table 4D: Proc Panel output of the fixed firm effects model

The PROC PANEL output in SAS reports parameter estimates and one specification test for fixed effects. The specification test is based on the F statistic for the hypothesis that all fixed effects parameters are equal to 0. The F statistic is computed as $\beta_f S_f^{-1} \beta_f / n$ with degrees of freedom equal to M – K. β_f is the estimated n dimensional vector of fixed-effects parameters, and S_f^{-1} is the estimated covariance matrix of the fixed-effects parameters. M is the total number of observations, and K is the number of explanatory variables (SAS Institute Inc. 2008). The total number of cross-sectional units is 862; therefore 861 dummy parameters and the intercept are estimated. The p-value of the F test is <.0001. It allows rejecting the null hypothesis at all practical levels of significance. The T tests for the effects of explanatory variables indicate that size, capital ratio, return on capital, health product risk, annuity product risk, and regulatory asset risk have significant relationships with the RBC ratio.

We estimate the two-way fixed effects model by adding time effect dummy variables to the model.

$LogRBCratio_{it} = \beta_0 + \beta_1*LogATotal_{it} + \beta_2*CAP_{it} + \beta_3*RetOnCap_{it} + \beta_4*ProdARisk_{it} + \beta_5*ProdHRisk_{it} + \beta_6*Preinsur_{it} + \beta_7*pRegARisk_{it} + \beta_8*LogpOppARisk_{it} + \alpha_i + \tau_t + \varepsilon_{it}$, where the α_i and the τ_t are dummy variables for the firm and time effects, respectively. Since there are 862 cross-sections and 4 time periods, the number of dummy parameters becomes 864. Dummy variables for the last insurer and the last time period are dropped from the model.

Model Description	
Estimation Method	FixTwo
Number of Cross Sections	862
Time Series Length	4

Table 5A

Fit Statistics			
SSE	205.7791	**DFE**	2353
MSE	0.0875	**Root MSE**	0.2957
R-Square	0.9164		

Table 5B

F Test for No Fixed Effects			
Num DF	Den DF	F Value	Pr > F
864	2353	16.80	<.0001

Table 5C

Variable	DF	Estimate	Standard Error	t Value	Pr > |t|	Label
LogATotal	1	0.201359	0.0264	7.64	<.0001	log(Total assets)
CAP	1	2.677488	0.0925	28.94	<.0001	Mkt capital / Total assets
RetOnCap	1	0.046104	0.00625	7.38	<.0001	Income / Mkt capital
ProdHRisk	1	-0.37818	0.0981	-3.86	0.0001	Health writings / Total writings
ProdARisk	1	-0.30369	0.1121	-2.71	0.0068	Annuity writings / Total writings
Preinsur	1	0.014527	0.0614	0.24	0.8131	Reinsurance writings / Total writings
pRegARisk	1	-9.025	0.5542	-16.28	<.0001	Regulatory asset risk / Total invested assets
pOppARisk	1	5.733557	2.1136	2.71	0.0067	Opportunity asset risk / Total invested assets
Ngroup	1	0.001921	0.0630	0.03	0.9757	Indicator for member of group (1=yes)
Ntype	1	0.137853	0.2424	0.57	0.5697	Org type (1=stock)

Table 5D: Proc Panel output of the fixed firm and time effects model

The F test for fixed effects is highly significant, and the null hypothesis of "no fixed effects" is rejected. Significance of the insurer effects is not uniform. However, the time effects are uniformly significant, but they are not significantly different from each other. Explanatory variables with significant effects in the model are the log of total assets, capital ratio, return on capital, health product risk, annuity product risk, regulatory asset risk, and opportunity asset risk. The presence of a great number of parameters inflates the R-Square up to 0.91. If we compare it to the two-way fixed-effects model, the estimates of the regression coefficients and the standard errors are almost identical. Only the coefficient estimate for opportunity asset risk becomes insignificant.

3.4. A random-effects model

To further investigate the effects of the explanatory variables, we consider a two-way random effects model.

$LogRBCratio_{it} = \beta_0 + \beta_1*LogATotal_{it} + \beta_2*CAP_{it} + \beta_3*RetOnCap_{it} + \beta_4*ProdARisk_{it} + \beta_5*ProdHRisk_{it} + \beta_6*Preinsur_{it} + \beta_7*pRegARisk_{it} + \beta_8*pOppARisk_{it} + Ngroup_{it} + Ntype_{it} + \alpha_i + \tau_t + \varepsilon_{it}$, where α_i is an insurer random effect, and the effects are iid from the distribution with the mean 0 and the variance σ_α^2. Likewise, the τ_t, time random effects, have mean 0 and the variance σ_τ^2. The random effects model allows including the time-invariant explanatory variables in a regression. The output from PROC PANEL gives estimates of variance components for cross-sections and time series.

In order to decide whether to use the fixed-effects or the random-effects model we can refer to the Hausman test results. A testing procedure suggested by Hausman (1978) notes that under H_0, the generalized least squares (GLS) for a random-effects model achieves the Cramer-Rao lower bounds, but under H_1, the GLS is a biased estimator. In contrast, coefficient of variation of β is consistent under both H_0 and H_1 (Baltagi 1995, 68-72). Simply, the null hypothesis states the random-effects model is preferred over the fixed-effects model.

Model Description	
Estimation Method	RanTwo
Number of Cross Sections	862
Time Series Length	4

Table 6A

Fit Statistics			
SSE	3027.9912	**DFE**	3217
MSE	0.9412	**Root MSE**	0.9702
R-Square	0.3254		

Table 6B

Variance Component Estimates	
Variance Component for Cross Sections	0.564134
Variance Component for Time Series	0.002383
Variance Component for Error	0.087454

Table 6C

Hausman Test for Random Effects		
DF	**m Value**	**Pr > m**
10	80.45	<.0001

Table 6D

Parameter Estimates						
Variable	**DF**	**Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**Label**
Intercept	1	4.906161	0.2326	21.09	<.0001	Intercept
LogATotal	1	0.05298	0.0106	4.98	<.0001	log(Total assets)
CAP	1	2.419074	0.0734	32.96	<.0001	Mkt capital / Total assets
RetOnCap	1	0.047596	0.00592	8.04	<.0001	Income / Mkt capital
ProdHRisk	1	-0.69032	0.0611	-11.30	<.0001	Health writings / Total writings
ProdARisk	1	-0.09118	0.0773	-1.18	0.2385	Annuity writings / Total writings
Preinsur	1	-0.07928	0.0489	-1.62	0.1053	Reinsurance writings / Total writings
pRegARisk	1	-7.92	0.4269	-18.55	<.0001	Regulatory asset risk / Total invested assets
pOppARisk	1	4.593034	2.0220	2.27	0.0232	Opportunity asset risk / Total invested assets
Ngroup	1	-0.02184	0.0449	-0.49	0.6265	Indicator for member of group (1=yes)
Ntype	1	0.057054	0.0895	0.64	0.5241	Org type (1=stock)

Table 6E: Proc Panel output of the random firm and time effects model

The Hausman m statistic is reported in the PROC PANEL output. Given β_a, a vector of the ordinary least squares estimates, and β_b, a vector of the generalized least squares, $m = (\beta_b - \beta_a)'(S_b - S_a)^{-1}(\beta_b - \beta_a)$. S_b and S_a are consistent estimates of the asymptotic covariance matrices. Then m statistic is distributed χ^2 with K degrees of freedom, and K is the number of the explanatory variables (SAS Institute Inc. 2008). The p-value of <.0001 allows rejecting the null hypothesis of no correlation between effects and explanatory variables.

3.5. Heteroscedasticity and autocorrelation robust standard errors estimation

A shortcoming of the models considered so far is accepting the OLS assumptions such as independence of errors across insurers and years. Cross-sectional dependence implies residuals of a given year may be correlated across different insurers. Likewise, time-series dependence of residuals may be present for a given firm. When the residuals are correlated across observations, OLS standard errors can be biased and either over or underestimate the true variability of the coefficient estimates (Petersen 2009). In our panel data, there may possibly be dependence between insurers that are members of a group of affiliated companies. Insurers that are members of an affiliated group of firms may have superior access to investment opportunities and may have different mechanisms for monitoring and/or controlling managerial performance and structuring their capital and asset risk (Baranoff and Sager 2008). Unfortunately, accounting for insurer and time effects by including dummy variables may not fully mitigate cross-sectional and time dependence of the residuals. Contrary to the time series data in which the time label gives a natural ordering and structure, general forms of dependence for cross-sectional dimension are difficult to formulate (Hsiao 2006). In this report we assume that there is no cross-sectional dependence and perform the time clustering to account for autocorrelation within individual insurers.

Additionally, the assumption of homoscedasticity may not hold, because the residuals of large insurers tend to be larger than the residuals for small insurers. We implemented the log transformation of the response variable, but the residuals of individual insurers still may be heteroscedastic.

PANEL procedure in SAS provides options for the time clustering while fitting the two-way fixed-effects model. For example, we can obtain heteroscedasticity and serial correlation consistent standard error estimators by Arellano (1987) for the parameters estimates of the least squares dummy variables model. Arellano (1987) considers a separate covariance matrix with unknown form for each unit in panel data. In Proc Panel, the covariance matrices are obtained from separate regressions for units, where the intercepts correspond to the coefficients of the dummy variables for the units in the least squares dummy variable regression. Let's denote the covariance matrix for *i*-th unit as Ω_i. Arellano (1987) uses $Var(\beta) = (X'X)^{-1}X'\Omega X(X'X)^{-1}$ as the variance expression for the parameter estimator. The middle term, $X'\Omega X$, is computed as the sum of $X_i'\Omega_i X_i$ for each unit. Another advantage of Arellano estimators is an assumption of large N, units, and small T, time periods, which is appropriate for our data.

The two-way fixed-effects model with time clustering reports identical coefficient estimates as the model without clustering. The same explanatory variables have statistically significant relationships with the RBC ratio at the 0.05 significance level. However, the t-values for health product risk, annuity product risk, and opportunity asset risk are not highly significant in comparison with OLS estimates. This is due to increased values of the coefficient standard errors.

Model Description	
Estimation Method	FixTwo
Number of Cross Sections	862
Time Series Length	4
Hetero. Corr. Cov. Matrix Estimator	4

Table 7A

Fit Statistics			
SSE	205.7791	**DFE**	2353
MSE	0.0875	**Root MSE**	0.2957
R-Square	0.9164		

Table 7B

F Test for No Fixed Effects			
Num DF	**Den DF**	**F Value**	**Pr > F**
864	2353	16.80	<.0001

Table 7C

LogATotal	1	0.201359	0.0745	2.70	0.0069	log(Total assets)
CAP	1	2.677488	0.2840	9.43	<.0001	Mkt capital / Total assets
RetOnCap	1	0.046104	0.00830	5.55	<.0001	Income / Mkt capital
ProdHRisk	1	-0.37818	0.1890	-2.00	0.0455	Health writings / Total writings
ProdARisk	1	-0.30369	0.1437	-2.11	0.0347	Annuity writings / Total writings
Preinsur	1	0.014527	0.1166	0.12	0.9008	Reinsurance writings / Total writings
pRegARisk	1	-9.025	0.9518	-9.48	<.0001	Regulatory asset risk / Total invested assets
pOppARisk	1	5.733557	2.5266	2.27	0.0233	Opportunity asset risk / Total invested assets
Ngroup	1	0.001921	0.1141	0.02	0.9866	Indicator for member of group (1=yes)
Ntype	1	0.137853	0.1406	0.98	0.3269	Org type (1=stock)

Table 7D: Proc Panel output of the fixed firm and time effects model

We find a difference between the estimation methods in terms of the estimated standard errors. For the Arellano estimation, the estimates are approximately two times larger for the size, capital ratio, health product risk, reinsurance, regulatory asset risk, and group member dummy variables. On contrary, the estimates of standard errors increased by a small amount for return on the capital and opportunity asset risk variables and decreased for the organization type dummy variable.

Variables	Estimates of coefficients	Exponentiated estimates of coefficients	Estimates of standard errors	
			OLS	Arellano
LogATotal	0.2014	1.2231	0.0264	0.0745
CAP	2.6775	14.5487	0.0925	0.284
RetOnCap	0.0461	1.0472	0.0063	0.0083
ProdHRisk	-0.3782	0.6851	0.0981	0.189
ProdARisk	-0.3037	0.7381	0.1121	0.1437
Preinsur	0.0145	1.0146	0.0614	0.1166
pRegARisk	-9.025	0.00012	0.5542	0.9518
pOppARisk	5.7336	309.08	2.1136	2.5266
Ngroup	0.0019	1.0019	0.063	0.1141
Ntype	0.1379	1.1479	0.2424	0.1406

Table 8: Comparison of two estimation methods for the firm and time fixed effects model

The exponential of a coefficient estimate yields the expected change in the RBC ratio for one unit change in a variable when the other variables are held constant.

Conclusion

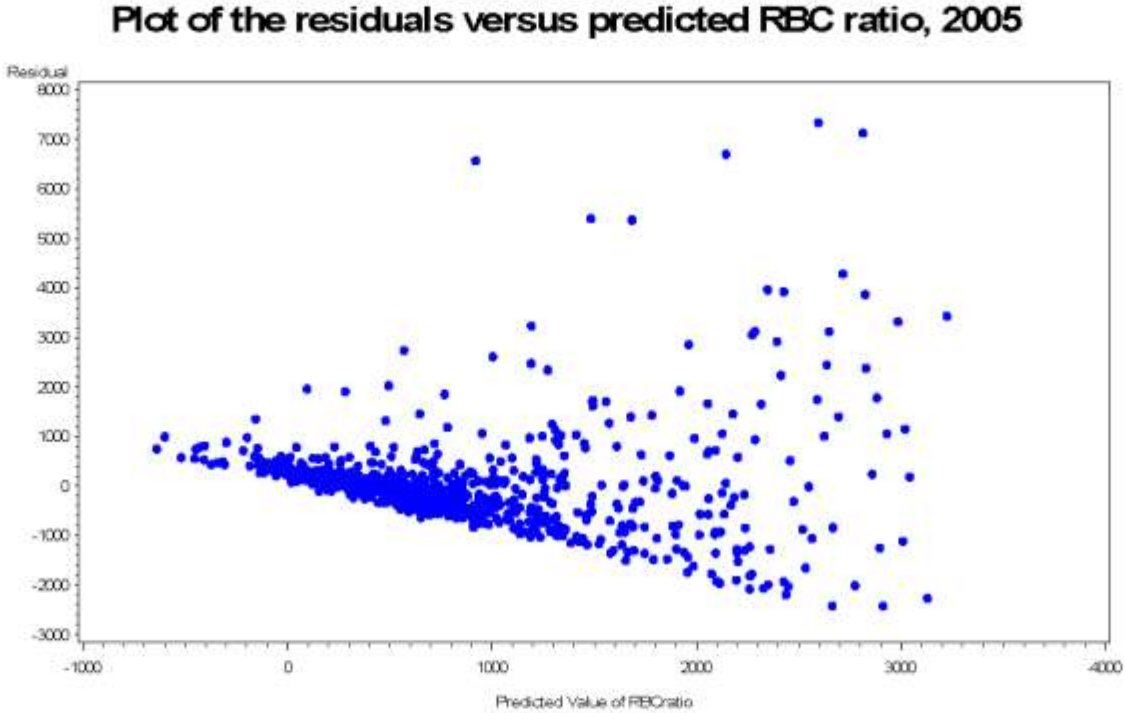
In this study we utilize panel data of life insurers during the period of 2005 to 2008 to discover variables that have statistically significant relationships with the RBC ratio. The use of panel data analysis is motivated by its ability to account for unobserved heterogeneity among the insurers and insurer-invariant time effects. Exploratory data analysis is conducted to check the error assumptions within each given year. The natural logarithm transformation of the response variable is performed to remedy the error heteroscedasticity across the insurers. The limitation of our study is that we assume no correlation between the insurers. We find the two-way fixed effects model has advantages over the model with only fixed firm effects. The Hausman test statistic indicates correlation between the fixed effects and the explanatory variables, thus we prefer the fixed effects model. First, the models are estimated based on the assumption of constant error variance and no autocorrelation for individual insurers. However, these assumptions probably do not hold. The Arellano method allows us to obtain the coefficient standard errors robust to autocorrelation and heteroscedasticity. The results suggest that the RBC ratio has a statistically significant relationship with size and capital structure of the insurers and the relationships are positive. The increase in return on capital and opportunity asset risk is associated with the increase in the RBC ratio. Health product risk, annuity product risk, and regulatory asset risk have a negative relationship with the RBC ratio. Further study may focus on causal effects of the explanatory variables on the RBC ratio of life insurers.

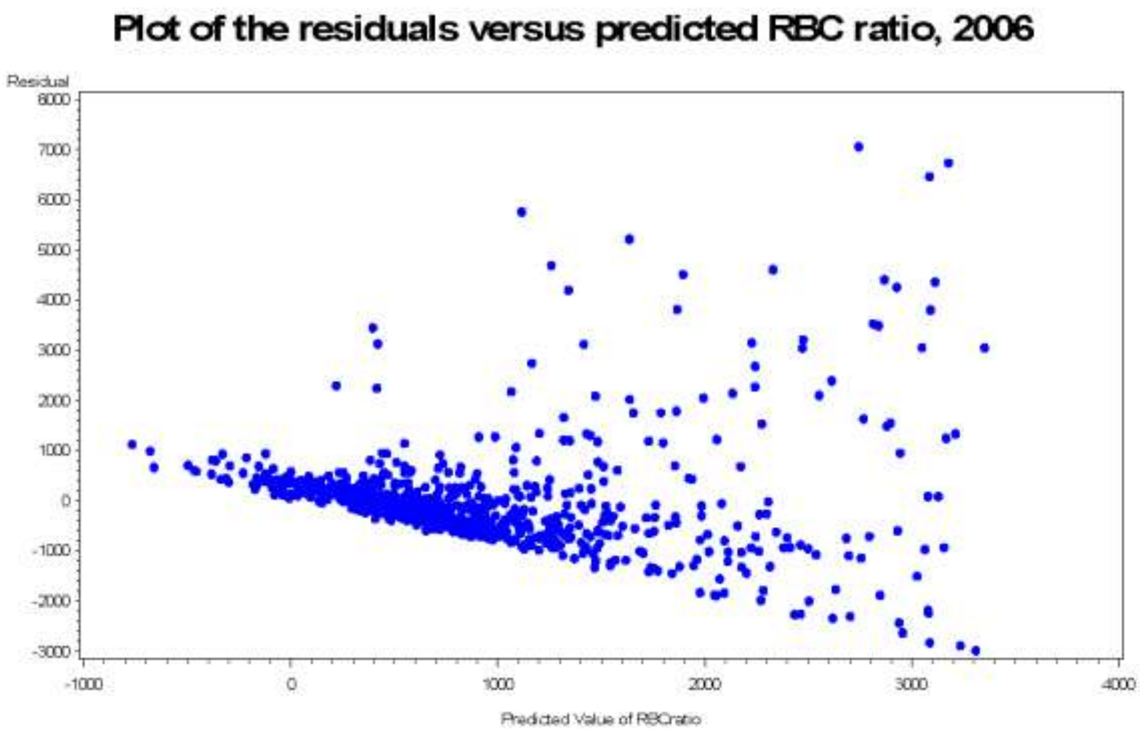
Appendix A: The Pearson correlation matrix

Pearson Correlation Coefficients Prob > |r| under H0: Rho=0 Number of Observations						
	Preinsur	**Ntype**	**Ngroup**	**ProdHRisk**	**ProdARisk**	**RBCratio**
Preinsur Reinsurance writings / Total writings	1.00000 3297	0.07075 <.0001 3297	-0.03943 0.0236 3297	-0.24815 <.0001 3297	-0.22941 <.0001 3297	0.08605 <.0001 3297
Ntype Org type (1=stock)	0.07075 <.0001 3297	1.00000 3317	0.10329 <.0001 3317	-0.02275 0.1915 3297	0.00524 0.7635 3297	0.04256 0.0142 3317
Ngroup Indicator for member of group (1=yes)	-0.03943 0.0236 3297	0.10329 <.0001 3317	1.00000 3317	0.02726 0.1176 3297	0.13128 <.0001 3297	-0.04811 0.0056 3317
ProdHRisk Health writings / Total writings	-0.24815 <.0001 3297	-0.02275 0.1915 3297	0.02726 0.1176 3297	1.00000 3297	-0.37672 <.0001 3297	-0.12018 <.0001 3297
ProdARisk Annuity writings / Total writings	-0.22941 <.0001 3297	0.00524 0.7635 3297	0.13128 <.0001 3297	-0.37672 <.0001 3297	1.00000 3297	-0.04880 0.0051 3297
RBCratio 100*Mkt cap / (2*Auth cap)	0.08605 <.0001 3297	0.04256 0.0142 3317	-0.04811 0.0056 3317	-0.12018 <.0001 3297	-0.04880 0.0051 3297	1.00000 3317
RetOnCap Income / Mkt capital	-0.00294 0.8659 3297	0.00072 0.9668 3317	0.00568 0.7437 3317	-0.00757 0.6638 3297	-0.02082 0.2320 3297	0.02356 0.1749 3317
CAP Mkt capital / Total assets	0.16404 <.0001 3297	0.04895 0.0048 3317	-0.07183 <.0001 3317	0.23383 <.0001 3297	-0.38396 <.0001 3297	0.45153 <.0001 3317
LogATotal log(Total assets)	-0.10156 <.0001 3297	-0.10257 <.0001 3317	0.34737 <.0001 3317	-0.18919 <.0001 3297	0.46553 <.0001 3297	-0.33305 <.0001 3317
pRegARisk Regulatory asset risk / Total invested assets	0.15088 <.0001 3297	-0.03119 0.0725 3317	-0.00493 0.7765 3317	-0.05270 0.0025 3297	-0.10987 <.0001 3297	-0.11001 <.0001 3317
pOppARisk Opportunity asset risk / Total invested assets	0.15656 <.0001 3297	0.02086 0.2297 3317	-0.00853 0.6233 3317	-0.00870 0.6176 3297	-0.12852 <.0001 3297	-0.00526 0.7620 3317

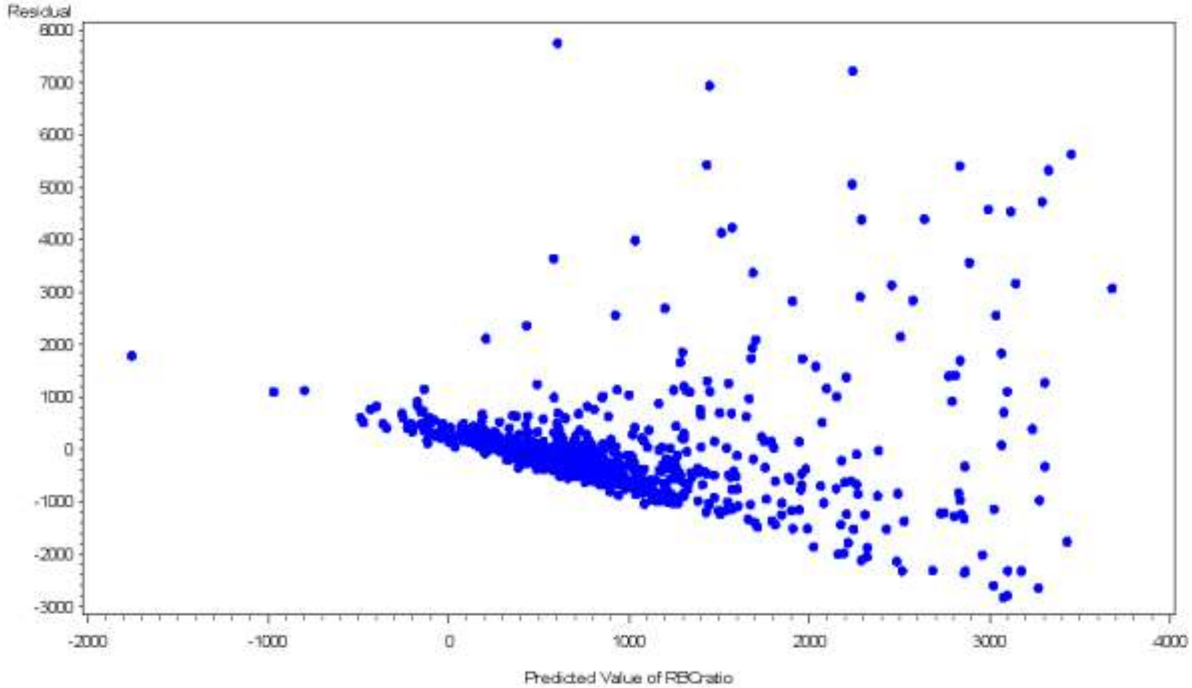
Pearson Correlation Coefficients Prob > |r| under H0: Rho=0 Number of Observations					
	RetOnCap	**CAP**	**LogATotal**	**pRegARisk**	**pOppARisk**
Preinsur Reinsurance writings / Total writings	-0.00294 0.8659 3297	0.16404 <.0001 3297	-0.10156 <.0001 3297	0.15088 <.0001 3297	0.15656 <.0001 3297
Ntype Org type (1=stock)	0.00072 0.9668 3317	0.04895 0.0048 3317	-0.10257 <.0001 3317	-0.03119 0.0725 3317	0.02086 0.2297 3317
Ngroup Indicator for member of group (1=yes)	0.00568 0.7437 3317	-0.07183 <.0001 3317	0.34737 <.0001 3317	-0.00493 0.7765 3317	-0.00853 0.6233 3317
ProdHRisk Health writings / Total writings	-0.00757 0.6638 3297	0.23383 <.0001 3297	-0.18919 <.0001 3297	-0.05270 0.0025 3297	-0.00870 0.6176 3297
ProdARisk Annuity writings / Total writings	-0.02082 0.2320 3297	-0.38396 <.0001 3297	0.46553 <.0001 3297	-0.10987 <.0001 3297	-0.12852 <.0001 3297
RBCratio 100*Mkt cap / (2*Auth cap)	0.02356 0.1749 3317	0.45153 <.0001 3317	-0.33305 <.0001 3317	-0.11001 <.0001 3317	-0.00526 0.7620 3317
RetOnCap Income / Mkt capital	1.00000 3317	0.04695 0.0068 3317	0.00259 0.8815 3317	-0.00093 0.9572 3317	-0.01010 0.5608 3317
CAP Mkt capital / Total assets	0.04695 0.0068 3317	1.00000 3317	-0.61252 <.0001 3317	0.30673 <.0001 3317	0.35584 <.0001 3317
LogATotal log(Total assets)	0.00259 0.8815 3317	-0.61252 <.0001 3317	1.00000 3317	-0.03034 0.0806 3317	-0.12863 <.0001 3317
pRegARisk Regulatory asset risk / Total invested assets	-0.00093 0.9572 3317	0.30673 <.0001 3317	-0.03034 0.0806 3317	1.00000 3317	0.75199 <.0001 3317
pOppARisk Opportunity asset risk / Total invested assets	-0.01010 0.5608 3317	0.35584 <.0001 3317	-0.12863 <.0001 3317	0.75199 <.0001 3317	1.00000 3317

Appendix B: Proc Reg residual plots versus fitted response variable for the multiple regressions (response is RBCratio)

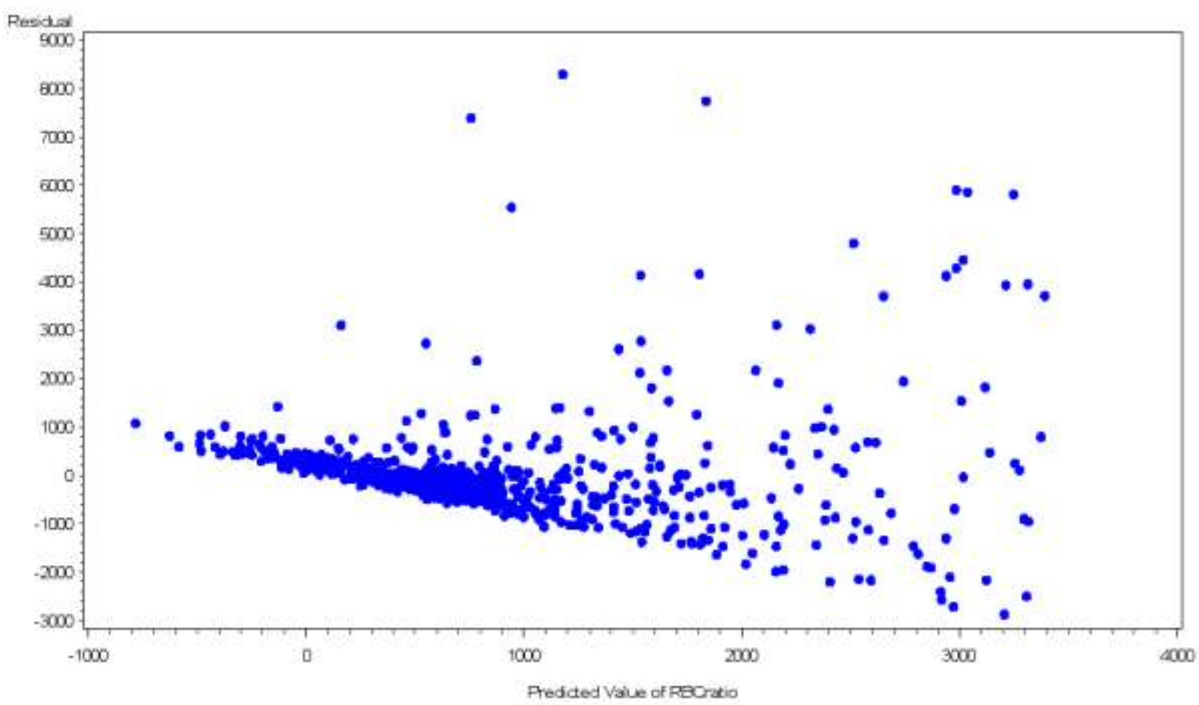




Plot of the residuals versus predicted RBC ratio, 2007

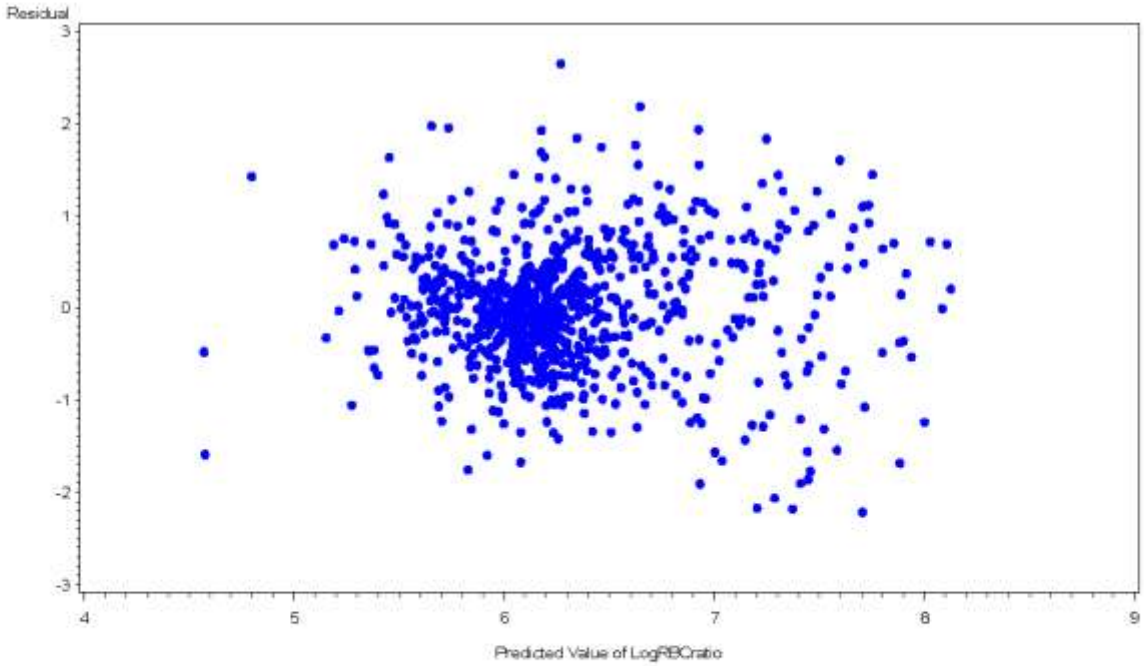


Plot of the residuals versus predicted RBC ratio, 2008

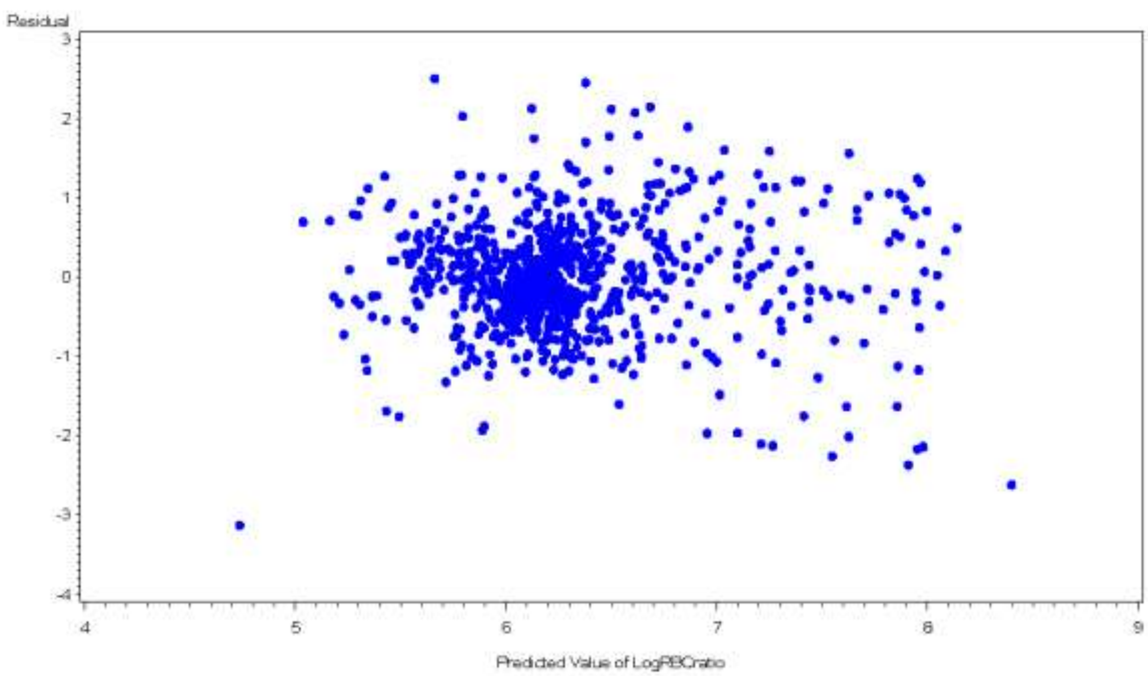


Appendix C: Proc Reg residual plots versus fitted response variable for the multiple regressions (response is LogRBCratio)

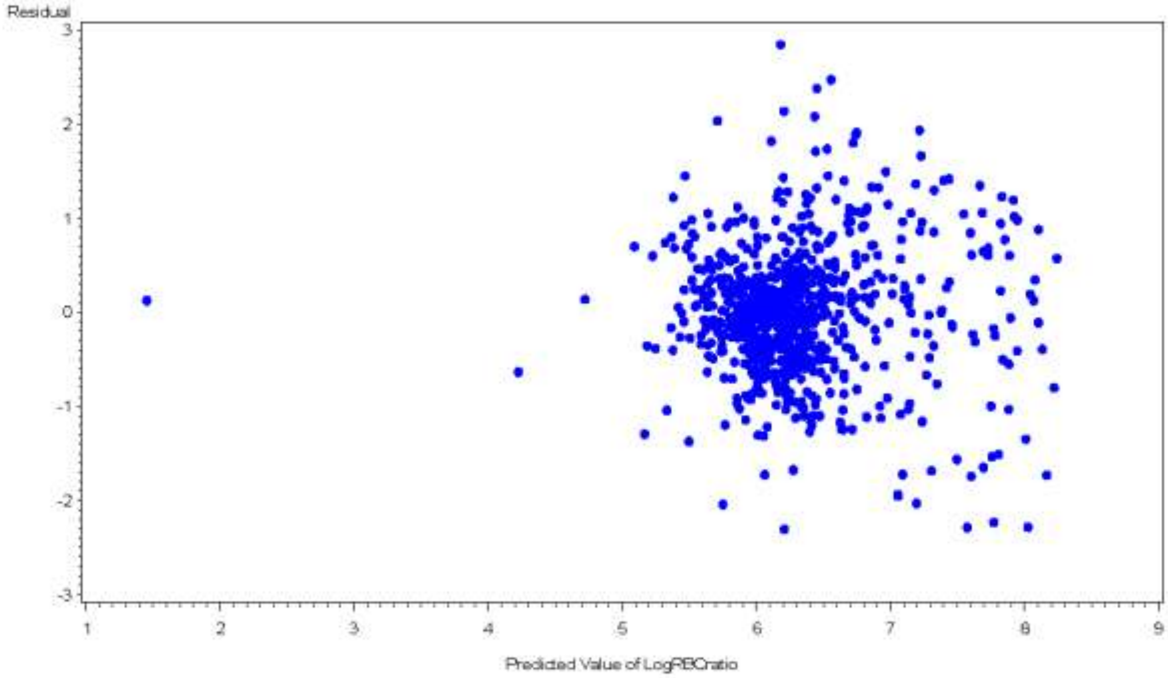
Plot of the residuals versus predicted LogRBCratio, 2005



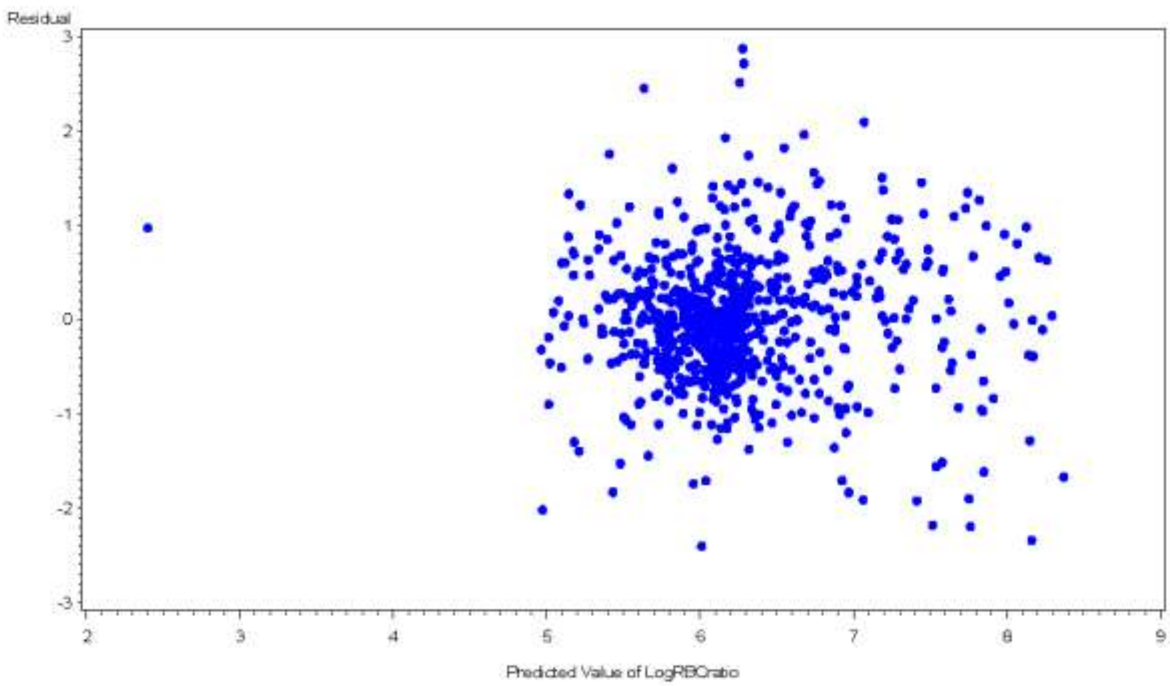
Plot of the residuals versus predicted LogRBCratio, 2006



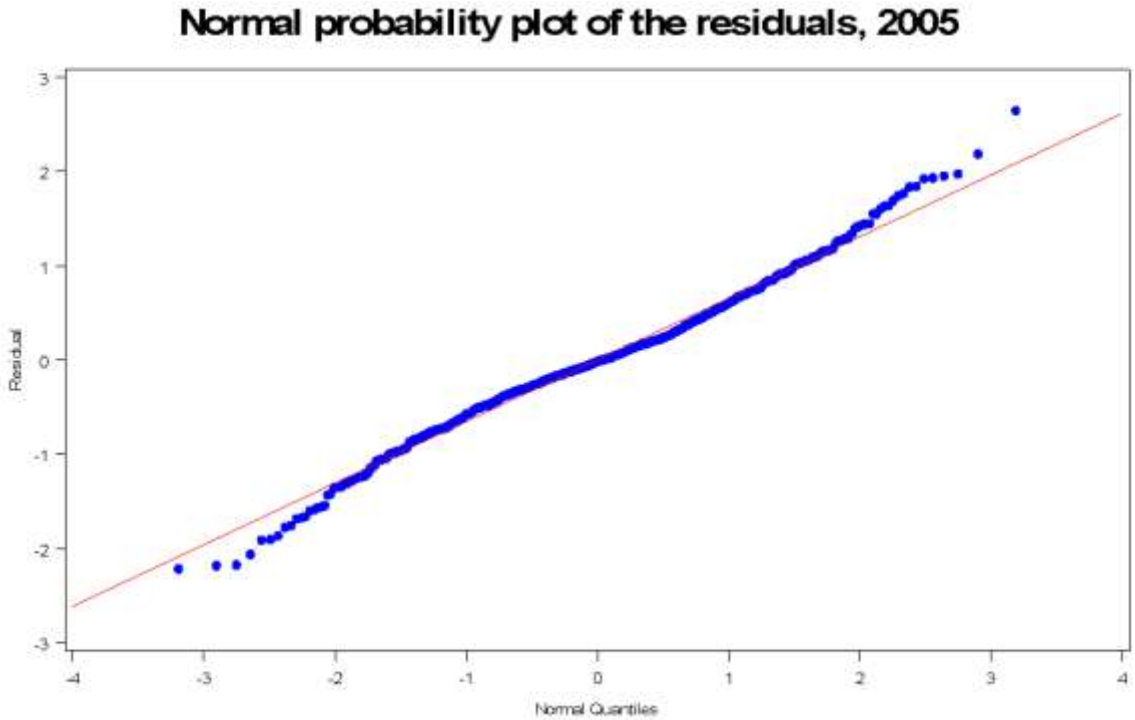
Plot of the residuals versus predicted LogRBCratio, 2007

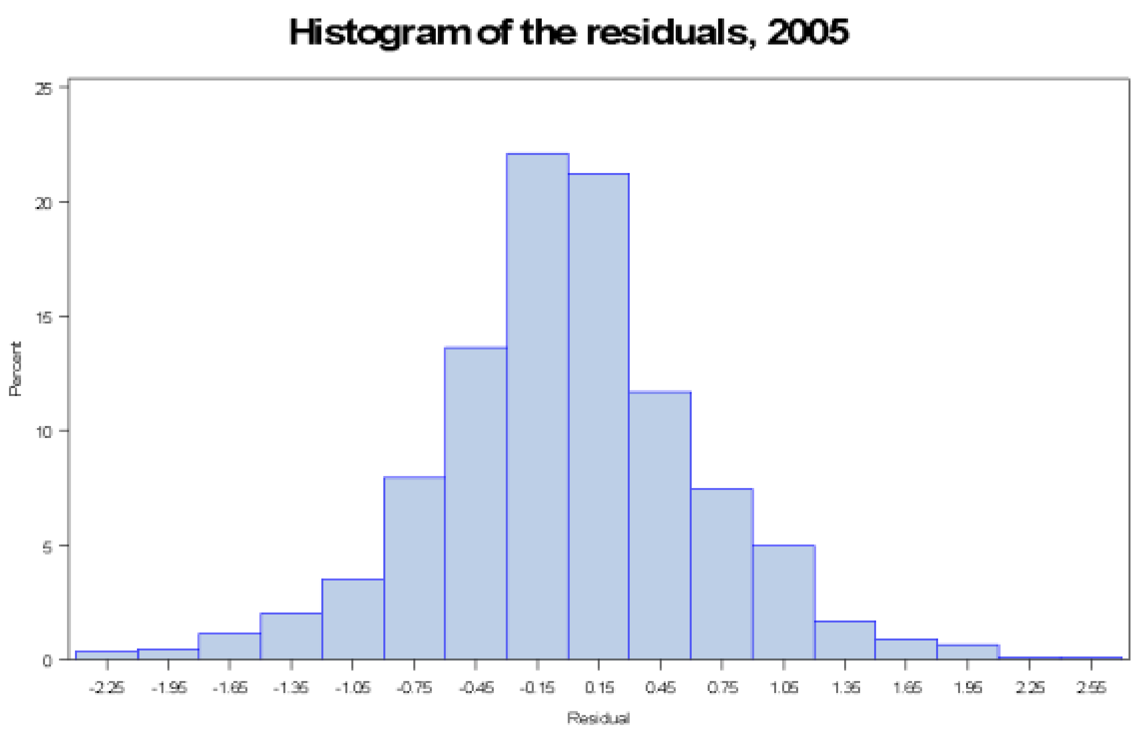


Plot of the residuals versus predicted LogRBCratio, 2008

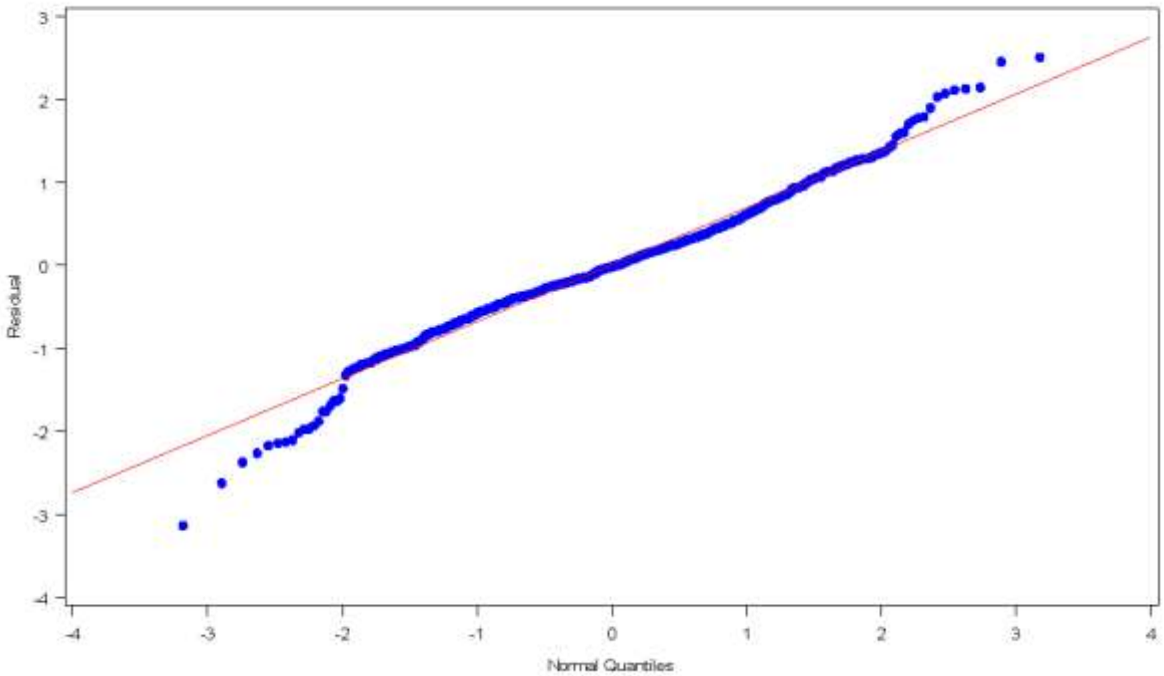


Appendix D: The histograms and the normal probability plots of the residuals (response is RBC ratio)

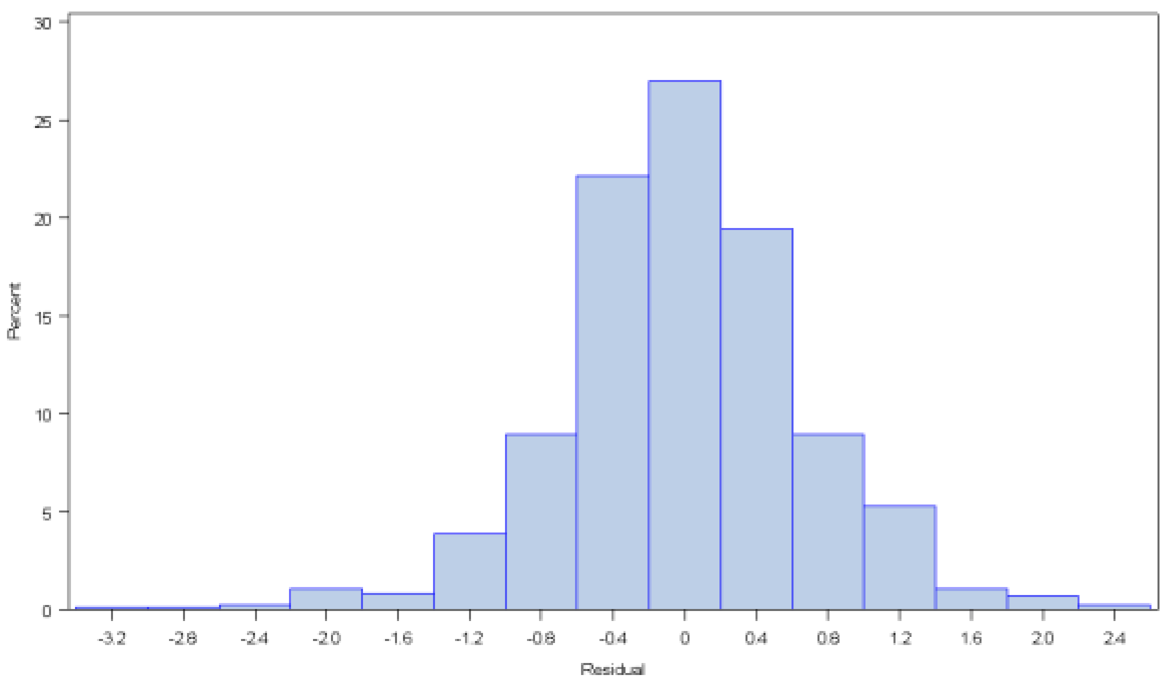




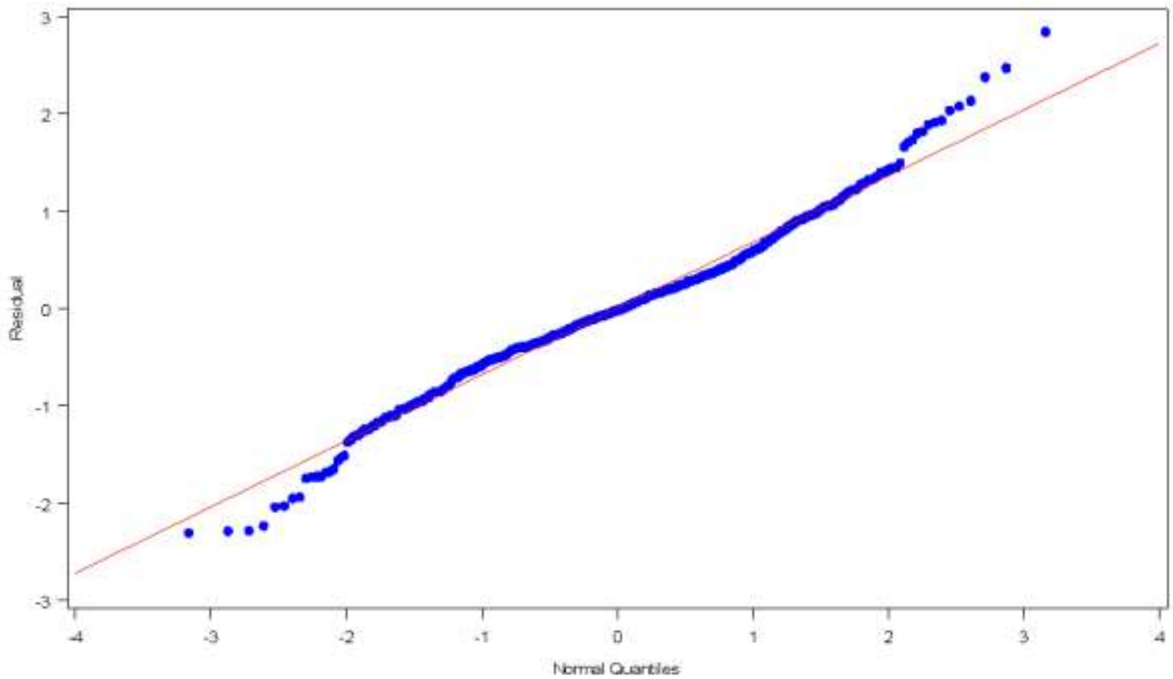
Normal probability plot of the residuals, 2006



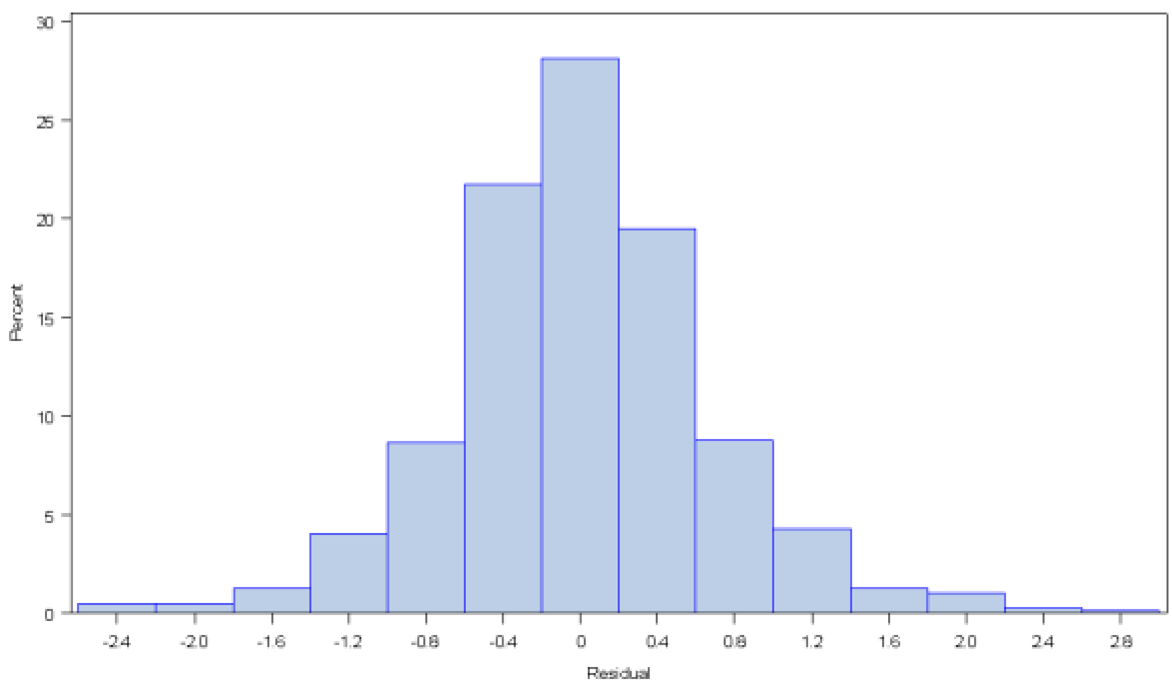
Histogram of the residuals, 2006



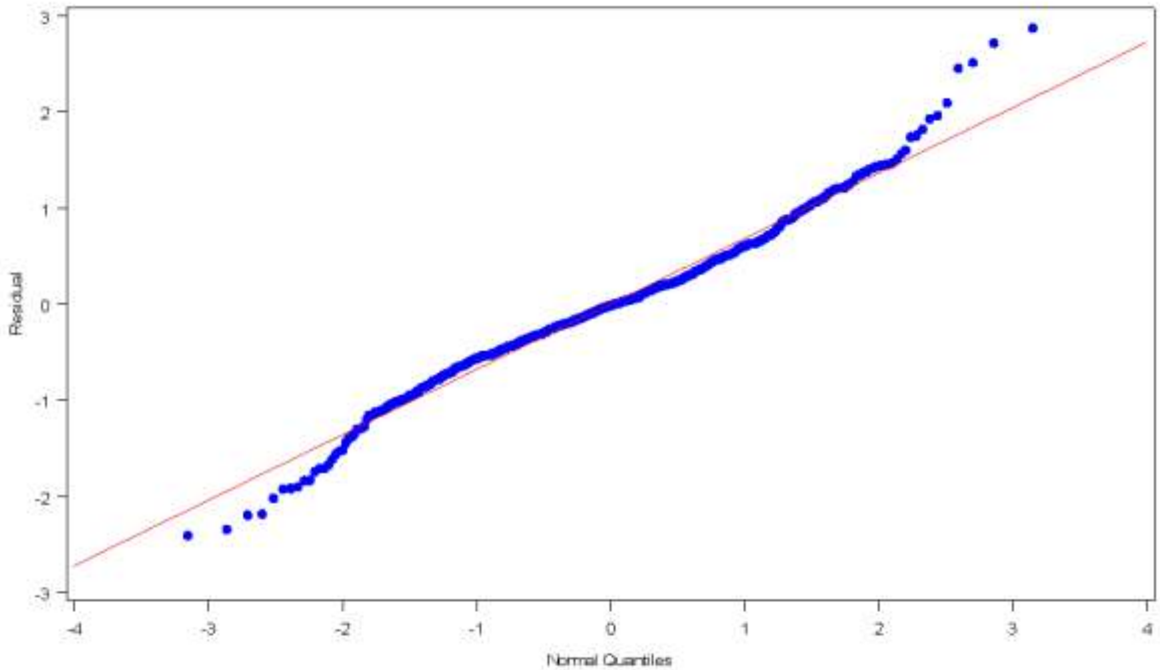
Normal probability plot of the residuals, 2007



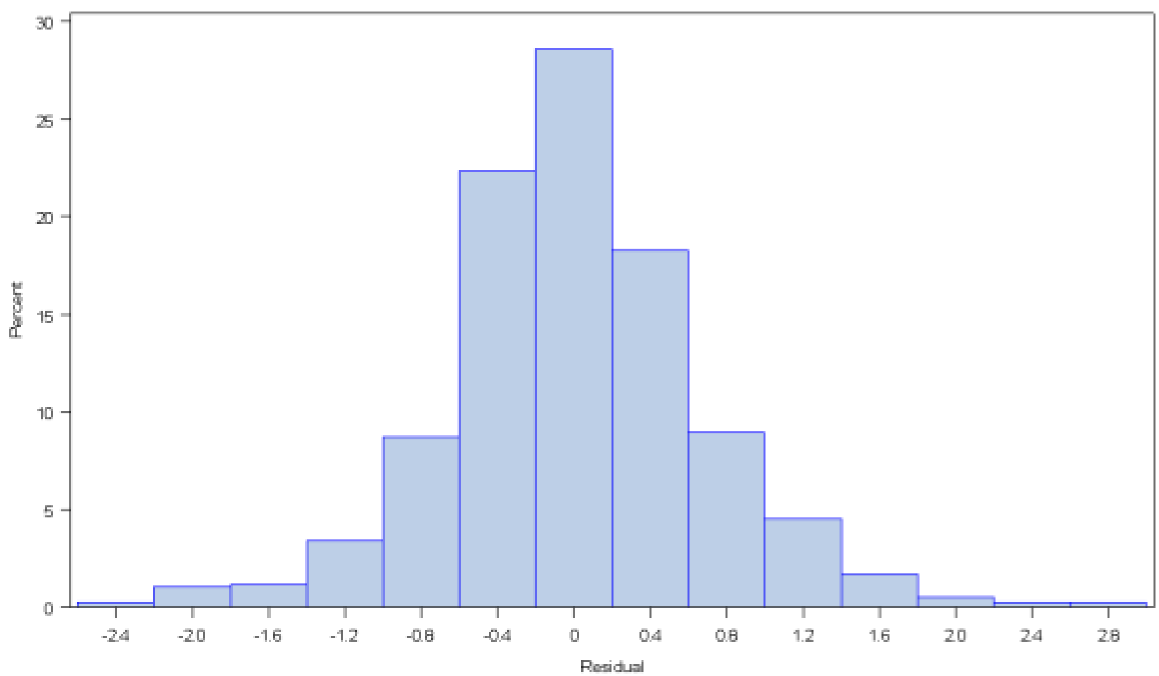
Histogram of the residuals, 2007



Normal probability plot of the residuals, 2008



Histogram of the residuals, 2008



References

Arellano, M. 1987. "Computing Robust Standard Errors for Within-groups Estimators." *Oxford Bulletin of Economics and Statistics* 49(4), 431-34.

Baltagi, B. 1995. *Econometric Analysis of Panel Data.* Chichester, England: John Wiley and Sons Ltd.

Baranoff, E. G., and Sager, T. W. 2002. "The relations among asset risk, product risk, and capital in the life insurance industry. *Journal of Banking and Finance* 26, 1181-97.

Baranoff, E. G., and Sager, T.W. 2008. Mortgage-backed securities and capital of life insurers: was the industry prepared for the credit crunch of 2007-2008? *Geneva Papers on Risk and Insurance – Issues and Practice, Special Competition Edition.* Retrieved from: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1215191

Baranoff, E. G., and Sager, T. W. 2011. The interplay between insurers' financial and asset risks during the crisis of 2007-2009. *Geneva Papers on Risk and Insurance – Issues and Practice* 36(3), 348-379.

Cummins, J. D., Harrington, S. E., and Klein R. 1995. Insolvency Experience, Risk-Based Capital, and Prompt Corrective Action in Property-Liability Insurance. *Journal of Banking and Finance* 19(3-4), 511-527.

Grace, M. F., Harrington, S. E., and Klein, R. W. 1998. Risk-Based Capital and Solvency Screening in Property-Liability Insurance: Hypotheses and Empirical Tests. *Journal of Risk and Insurance* 65(2), 213-243.

Greene, W. H. 2010. *Econometric Analysis. 7th edition.* Stern School of Business, New York University: Prentice Hall.

Hsiao, C. 1986. *Analysis of Panel Data.* New York, NY: The Press Syndicate of the University of Cambridge.

Hsiao, C. 2006. Panel Data Analysis – Advantages and Challenges. Institute of Economic Policy Research University of Southern California. Retrieved from http://dornsife.usc.edu/IEPR/Working%20Papers/IEPR_06.49_%5BHsiao%5D.pdf

National Association of Insurance Commissioners. 2009. Risk-Based Capital General Overview. Retrieved from http://www.naic.org/documents/committees_e_capad_RBCoverview.pdf

Petersen, M. A. 2009. "Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches." *Review of Financial Studies* 22(1), 435-480.

SAS Institute Inc. 2008. SAS/ETS 9.2 User's Guide. Cary, NC: SAS Institute Inc.